Conference Abstract

Persistent, Global, Unique: The three key requirements for a trusted identifier system for physical samples

Kerstin Lehnert[‡], Jens Klump[§], Lesley A Wyborn[|], Sarah Ramdeen[¶]

‡ Columbia University, New York, United States of America
§ CSIRO, Perth, Australia
| Australian National University, Canberra, Australia
¶ Ronin Institute, Huntsville, United States of America

Abstract

There is growing recognition that unambiguous citation and tracking of physical samples allows previously impossible linking of samples to data and publications, linking and integration of sample-based observations across data systems, and paves the road towards advanced data mining of sample-based data. And in recent years, there has been an uptake in the use of Persistent Identifiers (PIDs) for physical samples to support such citation and tracking.

The IGSN (International Geo Sample Number) is a PID for physical samples. It was originally developed for the solid earth sciences, and has evolved into an international PID system with members in five continents and a network of active allocating agents. It has been adopted by a growing number and range of stakeholders worldwide, including national geological surveys, research infrastructure providers, collection curators, researchers, and data managers, and by other disciplines that need to refer to physical samples. Nearly 6.9 million samples have been registered with IGSNs so far.

The IGSN system uses the Handle System (Kahn and Wilensky 1995; see also Handle.Net[®]) and has an international organization, IGSN e.V., to manage its governance structure and the technical architecture. The recent expansion of the IGSN beyond the geosciences into other domains such as biodiversity, archeology, and material sciences confirms the power of its concept and implementation, but imposes substantial pressures on the existing capacity and capabilities of the IGSN architecture and its governing organization. Modifications to the IGSN organizational and technical architecture are necessary at this point to keep pace with the growing demand and expectations. These changes are also necessary to ensure trustworthy and sustainable services for PID registration and resolution in a maturing research data ecosystem.

The essential criteria for a trustworthy system include an organizational foundation that ensures longevity, sustainability, proper governance, and regular quality assessment of registration services. It also includes a reliable and secure technical platform, based on open standards, which is sufficiently scalable and flexible to accommodate the growing diversity of specimen types, use cases, and stakeholder requirements.

In 2018, a major planning project for the IGSN was funded by the Alfred P. Sloan Foundation. An international group of experts participates in re-designing and improving the existing organization and technical architecture of the IGSN system, revising the current business model of the IGSN e.V. and professionalizing its operations. The goal is for the IGSN system to be able to respond to, and support in a sustainable manner, the rapidly growing demands of a global and increasingly multi-disciplinary user community, and to ensure that the IGSN will be a trustworthy, stable, and adaptable persistent identifier system for material samples, both technically and organizationally. The end result should also satisfy and facilitate participation across research domains, and will be a reliable component of the evolving research data ecosystem. Finally, it will ensure that the IGSN is recognized as a trusted partner by data infrastructure providers and the science community alike.

Keywords

identifier, physical sample, material sample, IGSN, PID

Presenting author

Kerstin Lehnert

Presented at

Biodiversity_Next 2019

References

- Kahn R, Wilensky R (1995) A Framework for Distributed Digital Object Services. Reston, VA: Corporation for National Research Initiatives Technical Note tn09-01 URL: http://hdl.handle.net/cnri.dlib/tn95-01